

amett F



SEQUENCE LISTING

KUMAGAI, Monto H.
DELLA-CIOPPA, Guy R.
ERWIN, Robert L.
McGEE, David R.

<120> METHOD OF COMPILING A FUNCTIONAL GENE PROFILE BY TRANSFECTING A
NUCLEIC ACID SEQUENCE OF A NON-PLANT DONOR INTO A HOST PLANT IN A
POSITIVE SENSE ORIENTATION

<130> 008010137US07

<140> 09/359,300

<141> 1999-07-21

<160> 71

<170> FastSEQ for Windows Version 3.0

<210> 1

<211> 26

<212> DNA

<213> Tomato mosaic virus

<400> 1

ctcgcaaagt ttcgaaccaa atcctc

26

<210> 2

<211> 35

<212> DNA

<213> Tomato mosaic virus

<400> 2

cggggtacct gggccccaac cggggggttc ggggg

35

<210> 3

<211> 41

<212> DNA

<213> Tomato mosaic virus

<400> 3

tctcgagcc taggctcgca aagtttcgaa ccaaatectc a

41

<210> 4

<211> 35

<212> DNA

<213> Tomato mosaic virus

<400> 4

cggggtacct gggccccaac cggggggttc ggggg

35

<210> 5

<210> 5
 <211> 24
 <212> DNA
 <213> Tomato mosaic virus

 <400> 5
 tatgtatggt gcagaagaac agat 24

 <210> 6
 <211> 24
 <212> DNA
 <213> Tomato mosaic virus

 <400> 6
 agtcgactct tcctcttctg gcat 24

 <210> 7
 <211> 30
 <212> DNA
 <213> Tomato mosaic virus

 <400> 7
 tgctcgagtg tgttcttcag tttctgtca 30

 <210> 8
 <211> 30
 <212> DNA
 <213> Tomato mosaic virus

 <400> 8
 aactcgagcg ctttgatttc tccgaagctt 30

 <210> 9
 <211> 114
 <212> DNA
 <213> Tomato mosaic virus

 <220>
 <221> CDS
 <222> (28)...(115)

 <400> 9
 gttttaaata cgctcgaggt tttaaata atg tct gtt gcc ttg tta tgg gtt gtt 54
 Met Ser Val Ala Leu Leu Trp Val Val
 1 5

 tct cct tgt gac gtc tca aat ggg aca agt ttc atg gaa tca gtc cgg 102
 Ser Pro Cys Asp Val Ser Asn Gly Thr Ser Phe Met Glu Ser Val Arg
 10 15 20 25

 gag gga aac cgt 114
 Glu Gly Asn Arg

<210> 10
<211> 29
<212> PRT
<213> Tomato mosaic virus

<400> 10
Met Ser Val Ala Leu Leu Trp Val Val Ser Pro Cys Asp Val Ser Asn
1 5 10 15
Gly Thr Ser Phe Met Glu Ser Val Arg Glu Gly Asn Arg
20 25

<210> 11
<211> 39
<212> DNA
<213> Nicotiana benthamiana

<400> 11
gcctcgagtg cagcatggaa acccttctaa agcttttcc 39

<210> 12
<211> 36
<212> DNA
<213> Nicotiana benthamiana

<400> 12
tccctaggtc aaaggtcttc tattgctaga ttgccc 36

<210> 13
<211> 111
<212> DNA
<213> Tobacco mosaic virus

<220>
<221> CDS
<222> (25)...(111)

<400> 13
gttttaaata cgctcgagtg cagc atg gaa acc ctt cta aag cct ttt cca 51
Met Glu Thr Leu Leu Lys Pro Phe Pro
1 5

tct cct tta ctt tcc att cct act cct aac atg tat agt ttc aaa cac 99
Ser Pro Leu Leu Ser Ile Pro Thr Pro Asn Met Tyr Ser Phe Lys His
10 15 20 25

aac ttc act ttt 111
Asn Phe Thr Phe

<210> 14
<211> 29
<212> PRT

<213> Tobacco mosaic virus

<400> 14

Met Glu Thr Leu Leu Lys Pro Phe Pro Ser Pro Leu Leu Ser Ile Pro
1 5 10 15
Thr Pro Asn Met Tyr Ser Phe Lys His Asn Phe Thr Phe
20 25

<210> 15

<211> 44

<212> DNA

<213> Erwinia herbicola

<400> 15

ccaagcttct cgagtgcagc atgcagcaac cgccgctgct tgac

44

<210> 16

<211> 43

<212> DNA

<213> Erwinia herbicola

<400> 16

aagatctctc gagctaaacg ggacgctgcc aaagaccggc cgc

43

<210> 17

<211> 23

<212> DNA

<213> Tobacco mild green mosaic virus

<400> 17

tgtgaaactc gaaaagggttc cgg

23

<210> 18

<211> 36

<212> DNA

<213> Tobacco mild green mosaic virus

<400> 18

cggggtacct gggccgctac cggcggttag gggagg

36

<210> 19

<211> 31

<212> DNA

<213> Ribgrass mosaic virus

<400> 19

tactcgaggt tcataagacc gcggtaggcg g

31

<210> 20

<211> 36

<212> DNA

<213> Ribgrass mosaic virus

<400> 20

cggggtacct gggcccctac ccgggggttta gggagg

36

<210> 21
<211> 107
<212> DNA
<213> N. tabacum

<220>
<221> CDS
<222> (21)...(107)

<400> 21
gttttaaata cgctcgagcc atg gct tcc tca gtt ctt tcc tct gca gca gtt 53
Met Ala Ser Ser Val Leu Ser Ser Ala Ala Val
1 5 10

gcc acc cgc agc aat gtt gct caa gct aac atg gtt gca cct ttc act 101
Ala Thr Arg Ser Asn Val Ala Gln Ala Asn Met Val Ala Pro Phe Thr
15 20 25

ggc ctt 107
Gly Leu

<210> 22
<211> 29
<212> PRT
<213> N. tabacum

<400> 22
Met Ala Ser Ser Val Leu Ser Ser Ala Ala Val Ala Thr Arg Ser Asn
1 5 10 15
Val Ala Gln Ala Asn Met Val Ala Pro Phe Thr Gly Leu
20 25

<210> 23
<211> 1543
<212> DNA
<213> Tobacco mild green mosaic virus

<400> 23
ctcgagggttc ataagaccgc ggtaggcgga gcgtttggtt actgtagtat aattaaatat 60
ttgtcagata aaagggttggt taaagatttg ttttttggtt gactgagtcg ataatgtctt 120
acgagcctaa agttagtgac ttccttgctc ttacgaaaaa ggaggaaatt ttaccaagg 180
ctttgacgag attaaagact gtctctatta gtactaagga tggtatatct gttaaggagt 240
ctgagtcctt gtgtgatatt gatttggttag tgaatgtgcc attagataag tatagggtatg 300
tgggtgtttt ggggtgttggt ttcaccgggtg aatggctggg accggatttc gttaaagggtg 360
gggtaacagt gagcgtgatt gacaaacggc ttgaaaattc cagagagtgc ataattggta 420
cgtaccgagc tgctgtaaaag gacagaagggt tccagttcaa gctgggtcca aattacttcg 480
tatccattgc ggatgccaaag cgaaaaccgt ggcagggttca tgtgcgaatt caaaatctga 540
agatcgaagc tggatggcaa cctctagctc tagagggtggg ttctgttgcc atgggttacta 600
ataacgtggg tggttaaagggt ttgagggaaa aggtcatcgc agtgaatgat ccgaacgtcg 660
aagggttcga aggtgtgggt gacgatttcg tcgattcggg tgctgcattc aaggcgattg 720

acagtttccg	aaagaaaaag	aaaaagattg	gaggaaggga	tgtaaataat	aataagtata	780
gatatagacc	ggagagatac	gccggtcctg	attcgttaca	atataaagaa	gaaaatgggt	840
tacaacatca	cgagctcgaa	tcagtaccag	tatttcgcag	cgatgtgggc	agagcccaca	900
gcgatgctta	accagtgcgt	gtctgcgttg	tcgcaatcgt	atcaaactca	ggcggcaaga	960
gatactgtta	gacagcagtt	ctctaacctt	ctgagtgcga	ttgtgacacc	gaaccagcgg	1020
tttccagaaa	caggataccg	ggtgtatatt	aattcagcag	ttctaaaacc	gttgtacgag	1080
tctctcatga	agtcctttga	tactagaaat	aggatcattg	aaactgaaga	agagtcgcgt	1140
ccatcggctt	ccgaagtatc	taatgcaaca	caacgtgttg	atgatgcgac	cgtggccatc	1200
aggagtcaaa	ttcagctttt	gctgaacgag	ctctccaacg	gacatgggtc	gatgaacagg	1260
gcagagttcg	aggtttttatt	accttgggct	actgcgccag	ctacataggc	gtggtgcaca	1320
cgatagtgca	tagtgttttt	ctctccactt	aaatcgaaga	gatatactta	cggtgtaatt	1380
ccgcaagggt	ggcgtaaacc	aaattacgca	atgttttagg	ttccatttaa	atcgaaacct	1440
gttatttcct	ggatcacctg	ttaacgtacg	cgtggcgat	attacagtgg	gaataactaa	1500
aagtgaagg	ttcgaatcct	ccctaacccc	gggtaggggc	cca		1543

<210> 24

<211> 55

<212> DNA

<213> rape mosaic virus

<400> 24

gatggcgcct	taatacgact	cactatagtt	ttatttttgt	tgcaacaaca	acaac	55
------------	------------	------------	------------	------------	-------	----

<210> 25

<211> 30

<212> DNA

<213> rape mosaic virus

<400> 25

cttgtgccct	tcatgacgag	ctatatcacg				30
------------	------------	------------	--	--	--	----

<210> 26

<211> 497

<212> DNA

<213> rape mosaic virus

<400> 26

ccttaatacg	actcactata	gtttttattt	tggttgcaaca	acaacaacaa	attacaataa	60
caacaaaaca	aatacaaaaca	acaacaacat	ggcacaattt	caacaaacag	taaacatgca	120
aacattgcag	gctgccgcag	ggcgcaacag	cctgggtgaat	gatttagcct	cacgacgtgt	180
ttatgacaat	gctgtcgagg	agctaaatgc	acgctcgaga	cgccctaagg	ttcattactc	240
caaatcagtg	tctacggaac	atgacgctgt	tagcttcaaa	cgtttatccg	gagtttgaga	300
tttcttttac	tcatacccaa	catgccgtac	actcccttgc	gggtggccta	aggactcttg	360
agttagagta	tctcatgatg	caagttccgt	tcggttctct	gacgtacgac	atcggtggta	420
actttgcagc	gcaccttttc	aaaggacgcg	actacgttca	ctgctgtatg	ccaaacttgg	480
atgtacgtga	tatagct					497

<210> 27

<211> 55

<212> DNA

<213> rape mosaic virus

<400> 27

gatggcgcct	taatacgact	cactatagtt	ttatttttgt	tgcaacaaca	acaac	55
------------	------------	------------	------------	------------	-------	----

<210> 28
 <211> 37
 <212> DNA
 <213> rape mosaic virus

 <400> 28
 atcgttttaa ctgggcccct acccggggtt agggagg 37

 <210> 29
 <211> 497
 <212> DNA
 <213> rape mosaic virus

 <400> 29
 ccttaatacg actcactata gttttatattt tgttgcaaca acaacaacaa attacaataa 60
 caacaaaaca aatacaaaaca acaacaacat ggcacaattt caacaaacag taaacatgca 120
 aacattccag gctgccgcag ggcgcaacag cctgggtgaat gatttagcct cagcagctgt 180
 ttatgacaat gctgtcgagg agctaaatgc acgctcgaga cgccctaagg ttcattactc 240
 caaatcagtg tctacggaac agacgctgtt agcttcaaac gcttatccgg agtttgagat 300
 ttcctttact catacccaaa catgccgtac actcccttgc gggtaggccta aggactcttg 360
 agttagagta tctcatgatg caagtccgt tgggttctct gacgtacgac atcgggtggta 420
 actttgcagc gcaccttttc aaaggacgag actacgttca ctgctgtatg ccaaacttgg 480
 atgtacgtga tatagct 497

 <210> 30
 <211> 55
 <212> DNA
 <213> rape mosaic virus

 <400> 30
 gatggcgct taatacgact cactatagtt ttatttttgt tgcaacaaca acaac 55

 <210> 31
 <211> 37
 <212> DNA
 <213> rape mosaic virus

 <400> 31
 atcgttttaa ctgggcccct acccggggtt agggagg 37

 <210> 32
 <211> 25
 <212> DNA
 <213> Picia pastoris

 <400> 32
 ttgcactctg ttggctcatg acgat 25

 <210> 33
 <211> 26
 <212> DNA
 <213> Pichia pastoris

<400> 33	
caagcttgca caaacgaacg tctcac	26
<210> 34	
<211> 42	
<212> DNA	
<213> <i>Pichia pastoris</i>	
<400> 34	
cactcgagag catggctatt cccgaagaat ttgatattat cg	42
<210> 35	
<211> 36	
<212> DNA	
<213> <i>Pichia pastoris</i>	
<400> 35	
tccttaggtt agaatctagc aagaccggtc ttctcg	36
<210> 36	
<211> 14	
<212> DNA	
<213> <i>Arabidopsis thaliana</i>	
<400> 36	
tcgagcggcc gcat	14
<210> 37	
<211> 42	
<212> DNA	
<213> <i>Trichosanthes kirilowii</i>	
<400> 37	
gcctcgagtg cagcatgatc agattcttag tcctctcttt gc	42
<210> 38	
<211> 36	
<212> DNA	
<213> <i>Trichosanthes kirilowii</i>	
<400> 38	
tccttaggtt aaatagcata acttcacat caaagc	36
<210> 39	
<211> 109	
<212> DNA	
<213> <i>Trichosanthes kirilowii</i>	
<220>	
<221> CDS	
<222> (20)...(109)	
<400> 39	
gttttaaata cgctcgagg atg atc aga ttc tta gtc ctc tct ttg cta att	52

	Met	Ile	Arg	Phe	Leu	Val	Leu	Ser	Leu	Leu	Ile	
	1				5					10		
ctc acc ctc ttc cta aca act cct gct gtg gag ggc gat gtt agc ttc												100
Leu Thr Leu Phe Leu Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe												
	15				20				25			
cgt tta tca												109
Arg Leu Ser												
	30											
<210>	40											
<211>	30											
<212>	PRT											
<213>	Trichosanthes kirilowii											
<400>	40											
Met Ile Arg Phe Leu Val Leu Ser Leu Leu Ile Leu Thr Leu Phe Leu												
1		5			10					15		
Thr Thr Pro Ala Val Glu Gly Asp Val Ser Phe Arg Leu Ser												
	20				25				30			
<210>	41											
<211>	19											
<212>	PRT											
<213>	P yoelii											
<400>	41											
Ser Tyr Val Pro Ser Ala Glu Gln Ile Leu Glu Phe Val Lys Gln Ile												
1		5			10					15		
Ser Ser Gln												
<210>	42											
<211>	839											
<212>	DNA											
<213>	Nicotiana benthamiana											
<220>												
<221>	CDS											
<222>	(15)...(677)											
<400>	42											
cttcactttc gccg atg gct cta cct aac cag caa acc gtg gat tac cct												50
	Met	Ala	Leu	Pro	Asn	Gln	Gln	Thr	Val	Asp	Tyr	Pro
	1				5					10		
agc ttc aag ctc gtt atc gtt ggc gat gga ggc aca ggg aag acc aca												98
Ser Phe Lys Leu Val Ile Val Gly Asp Gly Gly Thr Gly Lys Thr Thr												
	15				20				25			
ttt gta aag aga cat ctt act gga gag ttt gag aag aag tat gaa ccc												146
Phe Val Lys Arg His Leu Thr Gly Glu Phe Glu Lys Lys Tyr Glu Pro												

30	35	40	
act att ggt gtt gag gtt cat cct ctt gat ttc ttc act aac tgt ggc			194
Thr Ile Gly Val Glu Val His Pro Leu Asp Phe Phe Thr Asn Cys Gly			
45	50	55 60	
aag atc cgt ttc tac tgt tgg gat act gct ggc caa gag aaa ttt ggt			242
Lys Ile Arg Phe Tyr Cys Trp Asp Thr Ala Gly Gln Glu Lys Phe Gly			
65	70	75	
ggg ctt agg gat ggt tac tac atc cat gga caa tgt gct atc atc atg			290
Gly Leu Arg Asp Gly Tyr Tyr Ile His Gly Gln Cys Ala Ile Ile Met			
80	85	90	
ttt gat gtc aca gca cga ctg aca tac aag aat gtt cca aca tgg cac			338
Phe Asp Val Thr Ala Arg Leu Thr Tyr Lys Asn Val Pro Thr Trp His			
95	100	105	
cgt gat ctt tgc agg gtt tgt gaa aac atc cca att gtt ctt tgt ggg			386
Arg Asp Leu Cys Arg Val Cys Glu Asn Ile Pro Ile Val Leu Cys Gly			
110	115	120	
aat aaa gtt gat gtg aag aac agg caa gtc aag gcc aag cag gta aca			434
Asn Lys Val Asp Val Lys Asn Arg Gln Val Lys Ala Lys Gln Val Thr			
125	130	135 140	
ttc cac agg aag aag aac ctc cag tat tac gag ata tct gcc aag agc			482
Phe His Arg Lys Lys Asn Leu Gln Tyr Tyr Glu Ile Ser Ala Lys Ser			
145	150	155	
aac tac aac ttc gag aag cca ttc ttg tac ctt gct aga aag ctc gcc			530
Asn Tyr Asn Phe Glu Lys Pro Phe Leu Tyr Leu Ala Arg Lys Leu Ala			
160	165	170	
ggg gac gct aat ctt cac ttt gtg gaa tca cct gcc ctt gct ccc ccg			578
Gly Asp Ala Asn Leu His Phe Val Glu Ser Pro Ala Leu Ala Pro Pro			
175	180	185	
gaa gtt caa atc gac ttg gct gct cag cag cag cat gag gcg gag ctt			626
Glu Val Gln Ile Asp Leu Ala Ala Gln Gln Gln His Glu Ala Glu Leu			
190	195	200	
gca gca gca gca agt cag cca ctt cct gat gac gat gat gac acc ttc			674
Ala Ala Ala Ala Ser Gln Pro Leu Pro Asp Asp Asp Asp Thr Phe			
205	210	215 220	
gag tagagaaaga gagatgtgat ctgtcactga ttaccgcgtta gggcttgtct			727
Glu			
gaactttttt ttgttcatgg tgctattttt atgtgtccgt actttgaaat gaatcgatga			787
cattagtaat tttcattttt aagtttttaa ctgtcgctat gaaagtgaaa ac			839

<210> 43

<211> 221
 <212> PRT
 <213> Nicotiana benthamiana

<400> 43
 Met Ala Leu Pro Asn Gln Gln Thr Val Asp Tyr Pro Ser Phe Lys Leu
 1 5 10 15
 Val Ile Val Gly Asp Gly Gly Thr Gly Lys Thr Thr Phe Val Lys Arg
 20 25 30
 His Leu Thr Gly Glu Phe Glu Lys Lys Tyr Glu Pro Thr Ile Gly Val
 35 40 45
 Glu Val His Pro Leu Asp Phe Phe Thr Asn Cys Gly Lys Ile Arg Phe
 50 55 60
 Tyr Cys Trp Asp Thr Ala Gly Gln Glu Lys Phe Gly Gly Leu Arg Asp
 65 70 75 80
 Gly Tyr Tyr Ile His Gly Gln Cys Ala Ile Ile Met Phe Asp Val Thr
 85 90 95
 Ala Arg Leu Thr Tyr Lys Asn Val Pro Thr Trp His Arg Asp Leu Cys
 100 105 110
 Arg Val Cys Glu Asn Ile Pro Ile Val Leu Cys Gly Asn Lys Val Asp
 115 120 125
 Val Lys Asn Arg Gln Val Lys Ala Lys Gln Val Thr Phe His Arg Lys
 130 135 140
 Lys Asn Leu Gln Tyr Tyr Glu Ile Ser Ala Lys Ser Asn Tyr Asn Phe
 145 150 155 160
 Glu Lys Pro Phe Leu Tyr Leu Ala Arg Lys Leu Ala Gly Asp Ala Asn
 165 170 175
 Leu His Phe Val Glu Ser Pro Ala Leu Ala Pro Pro Glu Val Gln Ile
 180 185 190
 Asp Leu Ala Ala Gln Gln Gln His Glu Ala Glu Leu Ala Ala Ala
 195 200 205
 Ser Gln Pro Leu Pro Asp Asp Asp Asp Thr Phe Glu
 210 215 220

<210> 44
 <211> 738
 <212> DNA
 <213> Nicotiana benthamiana

<400> 44
 cttcactttc gccgatggct ctacctaacc agcaaacctg ggattaccct agcttcaagc 60
 tcgttatcgt tggcgatgga ggcacagggg agaccacatt tgtaaagaga catcttactg 120
 gagagtttga gaagaagtat gaaccacta ttggtgttga ggttcacact cttgatttct 180
 tcactaactg tggcaagatc cgtttctact gttggatact gctggccaag agaaatttgg 240
 tggctcttagg gatggttact acatccatgg acaatgtgct atcatcatgt ttgatgtcac 300
 agcacgactg acatacaaga atgttccaac atggcacctg gatctttgca gggttttgtg 360
 aaaacatccc aattgttctt tgtgggaata aagttgatgt gaagaacagg caagtcaagg 420
 ccaagcaggt aacattccac aggaagaaga acctccagta ttacgagata tctgccaga 480
 gcaactacaa cttcgagaag ccattcttgt accttgctag aaagctcgcc ggggacgcta 540
 atcttcactt tgtggaatca cctgcccttg ctcccccgga agttcaaate gacttggtg 600
 ctcagcagca gcatgaggcg gagcttgag cagcagcaag tcagccactt cctgatgacg 660
 atgatgacac cttcgagtag agaaagagag atgtgatctg tcaactgatta cccgtaggg 720
 cttgtctgaa cttttttt 738

<210> 45
 <211> 679
 <212> DNA
 <213> *Arabidopsis thaliana*

<400> 45
 cttcacttttc gccgatggct ctacctaacc agcaaaccgt ggattaccct agcttcaagc 60
 tcgttatcgt tggcgatgga ggcacaggga agaccacatt tgtaaagaga catcttactg 120
 gagagtttga gaagaagtat gaaccacta ttggtgttga ggttcatcct cctgatttct 180
 tcactaactg tggcaagatc cgtttctact gttgggatac tgctggccaa gagaaatttg 240
 gtggtcttag ggatggttac tacatccatg gacaatgtgc tatcatcatg tttgatgtca 300
 cagcagcact gacatacagg aatgttccaa catggcaccg tgatctttgc aggggtttgtg 360
 ccaagcaggt aacattccac aggaaggagg aactccagta ttacgagata tctgccaaga 420
 gcaactacaa cttcgagaag ccattcttgt accttgctag aaagctcgcg ggggacgcta 480
 atcttcactt tgtggaatca cctgcccttg ctcccccgga agttcaaate gacttggtcg 540
 ctcagcagca gcatgaggcg gagcttgca gacagcaag tcagccactt cctgatgacg 600
 atgatgacac cttcgagtag agaaagagag atgtgatctg tcaactgatta cccgttaggg 660
 cttgtctgaa cttttttttt 679

<210> 46
 <211> 667
 <212> DNA
 <213> *N. tabacum*

<400> 46
 atggctctac ctaaccagca aaccgtggat taccctagct tcaagctcgt tategttggc 60
 gatggaggca cagggaagac cacatttgta aagagacatc ttactggaga gtttgagaag 120
 aagtatgaac ccactattgg tgttgagggt catcctcttg atttcttcac taactgtggc 180
 aagatccgtt tctactgttg ggatactgct ggccaagaga aatttggtgg tcttagggat 240
 ggttactaca tccatggaca atgtgctatc atcatgtttg atgtcacagc acgactgaca 300
 tacaagaatg ttccaacatg gcaccgtgat ctttgccaggg tttgtgaaaa catcccaatt 360
 gttctttgtg ggaataaaat tgatgtgaag aacaggcaag tcaaggccaa gcaggtaaca 420
 ttccacagga agaagaacct ccagtattac gagatatctg ccaagagcaa ctacaacttc 480
 gagaagccat tcttgtaact tgctagaaaag ctgcgcgggg acgctaactt tcaactttgtg 540
 gaatcacctg cccttgctcc cccggaagtt caaatcgact tggctgctca gcagcagcat 600
 gaggcggagc ttgcagcagc agcaagtcag ccacttctctg atgacgatga tgacaccttc 660
 gtaga 667

<210> 47
 <211> 667
 <212> DNA
 <213> *N. tabacum*

<400> 47
 atggctctac caaaccaaca aactgtagat tatccaagct tcaagcttgt aatcgtgggc 60
 gatggaggaa ctgggaaaac aacttttgtc aagaggcatc ttactggtga atttgagaag 120
 aaatatgaac ccactattgg tgttgagggt catccattag acttcttcac aaattgtggg 180
 aaaattcgct tttattgtcg ggatactgct ggacaagaga agtttgaggg tcttcgggat 240
 ggttactaca ttcatgggca atgcgcaatt atcatgtttg atgttacagc ccgtctgacc 300
 tacaagaatg ttctactgtg gcacgcagat ctctgcaggg tttgtgaaaa catccccatt 360
 gttctttgtg gaaacaaaat tgatgtcaag aacaggcagg ttaaggcaaa gcaagttacc 420
 ttccacagga agaaaaattt gcaatactat gagatctcag caaagagtaa ctacaacttt 480
 gagaagcctt ttctgtacct tgccagaaaag cttgctgggg atgctaactt tcaactttgtg 540
 gaatcacctg cacttgctcc ccctgaagta caaattgatt tagctgcaca gcaactgcat 600

gaacaagagc ttttgcaagc cgctgcgcac gcacttccag atgacgatga tgaagctttt
gaataga

660
667

<210> 48
<211> 137
<212> PRT
<213> Tobacco mosaic virus

<400> 48
Met Ala Leu Pro Asn Gln Gln Thr Val Asp Tyr Pro Ser Phe Lys Leu
1 5 10 15
Val Ile Val Gly Asp Gly Gly Thr Gly Lys Thr Thr Phe Val Lys Arg
20 25 30
His Leu Thr Gly Glu Phe Glu Lys Lys Tyr Glu Pro Thr Ile Gly Val
35 40 45
Glu Val His Pro Leu Asp Phe Phe Thr Asn Cys Gly Lys Ile Arg Phe
50 55 60
Tyr Cys Trp Asp Thr Ala Gly Gln Glu Lys Phe Gly Gly Leu Arg Asp
65 70 75 80
Gly Tyr Tyr Ile His Gly Gln Cys Ala Ile Ile Met Phe Asp Val Thr
85 90 95
Ala Arg Leu Thr Tyr Lys Asn Val Pro Thr Trp His Arg Asp Leu Cys
100 105 110
Arg Val Cys Glu Asn Ile Pro Ile Val Leu Cys Gly Asn Lys Val Asp
115 120 125
Val Lys Asn Arg Gln Val Lys Ala Lys
130 135

<210> 49
<211> 135
<212> PRT
<213> Tobacco Mosaic Virus

<400> 49
Met Ala Leu Pro Asn Gln Gln Thr Val Asp Tyr Pro Ser Phe Lys Leu
1 5 10 15
Val Ile Val Gly Asp Gly Gly Thr Gly Lys Thr Thr Phe Val Lys Arg
20 25 30
His Leu Thr Gly Glu Phe Glu Lys Lys Tyr Glu Pro Thr Ile Gly Val
35 40 45
Glu Val His Pro Leu Asp Phe Phe Thr Asn Cys Gly Lys Ile Arg Phe
50 55 60
Tyr Cys Trp Asp Thr Ala Gly Gln Glu Lys Phe Gly Gly Leu Arg Asp
65 70 75 80
Gly Tyr Tyr Ile His Gly Gln Cys Ala Ile Ile Met Phe Asp Val Thr
85 90 95
Ser Thr Thr Asp Ile Gln Glu Cys Ser Asn Met Ala Pro Ser Leu Gln
100 105 110
Gly Leu Lys His Ser Gln Leu Phe Val Gly Ile Lys Leu Met Lys
115 120 125
Asn Arg Gln Val Lys Ala Gln
130 135

<210> 50

<211> 277
 <212> DNA
 <213> Tobacco mosaic virus

<220>
 <221> CDS
 <222> (1)...(277)

<400> 50
 gct act atg gtt gcc tct ccg gct cag gcc act atg gtc gct cct ttc 48
 Ala Thr Met Val Ala Ser Pro Ala Gln Ala Thr Met Val Ala Pro Phe
 1 5 10 15
 aac gga ctt aag tcc tcc gct cct tcc cag cca ccc gca agg cta aca 96
 Asn Gly Leu Lys Ser Ser Ala Pro Ser Gln Pro Pro Ala Arg Leu Thr
 20 25 30
 acg aca tta ctt cca tca caa gca acg gcg gaa gag tta act gca tgc 144
 Thr Thr Leu Leu Pro Ser Gln Ala Thr Ala Glu Glu Leu Thr Ala Cys
 35 40 45
 agg tgt ggc ctc cga ttg gaa aga aga agt ttg aga ctc tct ctt acc 192
 Arg Cys Gly Leu Arg Leu Glu Arg Arg Ser Leu Arg Leu Ser Leu Thr
 50 55 60
 ttc ctg acc tta ccg att ccg aat tgg cta agg aag ttg act acc tta 240
 Phe Leu Thr Leu Pro Ile Pro Asn Trp Leu Arg Lys Leu Thr Thr Leu
 65 70 75 80
 tcc gca aca agt gga ttc ctt gtg ttg aat tcg aag t 277
 Ser Ala Thr Ser Gly Phe Leu Val Leu Asn Ser Lys
 85 90

<210> 51
 <211> 92
 <212> PRT
 <213> Tobacco mosaic virus

<400> 51
 Ala Thr Met Val Ala Ser Pro Ala Gln Ala Thr Met Val Ala Pro Phe
 1 5 10 15
 Asn Gly Leu Lys Ser Ser Ala Pro Ser Gln Pro Pro Ala Arg Leu Thr
 20 25 30
 Thr Thr Leu Leu Pro Ser Gln Ala Thr Ala Glu Glu Leu Thr Ala Cys
 35 40 45
 Arg Cys Gly Leu Arg Leu Glu Arg Arg Ser Leu Arg Leu Ser Leu Thr
 50 55 60
 Phe Leu Thr Leu Pro Ile Pro Asn Trp Leu Arg Lys Leu Thr Thr Leu
 65 70 75 80
 Ser Ala Thr Ser Gly Phe Leu Val Leu Asn Ser Lys
 85 90

<210> 52

<211> 167
 <212> DNA
 <213> Arabidopsis thaliana

<400> 52
 acttgatctg tttcatacta aaacccaaaac tcatgtttgt tcaactccaaa cacaaacaca 60
 gcagtaatca aaaatcgtct tataacaaaa ggaaatgcaa caaacagaa gaaacaacta 120
 agtagtaggc aagattcttc ttcactcgtc ttcttggcta cggagcc 167

<210> 53
 <211> 393
 <212> DNA
 <213> Arabidopsis thaliana

<400> 53
 gaaacgacgt cggctagtta ttgggcatgg cctgaccagc agcaacaaca tcacaatcat 60
 catcagttca attgatcata ttgtctaaga acaacatcat actcatcttg atatcattat 120
 ttatcatcaa aagaaaattc cgtagatttt ttttaataagt attttcaa attttggcac 180
 gttaaattt aattaaattg gggtattatg tttacttgat ctgtttcata ctaaaaccaa 240
 aaggaaaacc aaaactcatg tttgttccact ccaaacacaa acacagcagt aatcaaaaat 300
 cgtcttataa caaaaaggaa atgcaacaaa acagaagaaa caactaagta gtaggcaaga 360
 ttctttcttca ctgctcttct tggctacgga gcc 393

<210> 54
 <211> 24
 <212> PRT
 <213> Arabidopsis thaliana

<400> 54
 Glu Thr Thr Ser Ala Ser Tyr Trp Ala Trp Pro Asp Gln Gln Gln Gln
 1 5 10 15
 His His Asn His His Gln Phe Asn
 20

<210> 55
 <211> 438
 <212> DNA
 <213> Homo sapiens

<400> 55
 gaagcggctc gccgcacag tgatgaagtg cgggaagggc aaagtttggc tcgatcccaa 60
 cgaaagctcc gacatctcca tggccaattc ccgccaaaac atcaggaagc ttgtgaagga 120
 tggtttcatc atcaggaagc caaccaagat tcaactctcgt tccagagctc gcaaaatgaa 180
 gattgccaag atgaagggtc gtcactctgg atacggtaag aggaagggta cccgtgaagc 240
 taggttgcca acaaagggtac tgtggatgag taggatgagc gttcttaggc gtctgttgaa 300
 gaaatacaga gagacgaaga agattgacaa gcacatgtac catgacatgt acatgcgtgt 360
 taagggtaat gtgttcaaga acaagcgtgt cttgatggag agtatccaca agtcaaaggc 420
 ttagaagcta ggggagaa 438

<210> 56
 <211> 438
 <212> DNA
 <213> Homo sapiens

<400> 56
gaagaggctc gcctctagtg tcttccgctg tggcaagaag aaggtctggt tagaccccaa 60
tgagaccaat gaaatcgcca atgccaactc ccgtcagcag atccggaagc tcatcaaaga 120
tgggctgata atccgcaagc ctgtgacggt ccattcccgg gctcgatgcc ggaaaaacac 180
cttgccccgc cggaaggga ggcacatggg cataggtaag cggaagggtta cagccaatgc 240
ccgaatgcca gagaagggtca catggatgag gagaatgagg attttgcgcc ggctgctcag 300
aagataccgt gaatctaaga agatcgatcg ccacatgtat cacagcctgt acctgaagggt 360
gaaggggaat gtgttcaaaa acaagcggat tctcatggaa cacatccaca agctgaaggc 420
agacaaggcc cgcaagaa 438

<210> 57
<211> 438
<212> DNA
<213> Homo sapiens

<400> 57
gaagcggctc gccgcacag tgatgaagtg cgggaaggga aaagtttggc tcgatcccaa 60
cgaaagctcc gacatctcca tggccaattc ccgcaaaaac atcaggaagc ttgtgaagga 120
tggtttcatc atcaggaagc caaccaagat tcaactctctg tccagagctc gcaaaatgaa 180
gattgccaag atgaagggtc gtcactctgg atacggtaag aggaagggtta cccgtgaagc 240
taggttgcca acaaagggtac tgtggatgag taggatgagt gttcttaggc gtctgttgaa 300
gaaatacaga gagacgaaga agattgacaa gcacatgtac catgacatgt acatgcgtgt 360
taagggtaat gtgttcaaga acaagcgtgt cttgatggag agtatccaca agtcaaaggc 420
ttagaagcta ggggagaa 438

<210> 58
<211> 438
<212> DNA
<213> Homo sapiens

<400> 58
gaagaggctc gcctctagtg tcttccgctg tggcaagaag aaggtctggt tagaccccaa 60
tgagaccaat gaaatcgcca atgccaactc ccgtcagcag atccggaagc tcatcaaaga 120
tgggctgata atccgcaagc ctgtgacggt ccattcccgg gctcgatgcc ggaaaaacac 180
cttgccccgc cggaaggga ggcacatggg cataggtaag cggaagggtta cagccaatgc 240
ccgaatgcca gagaagggtca catggatgag gagaatgagg attttgcgcc ggctgctcag 300
aagataccgt gaatctaaga agatcgatcg ccacatgtat cacagcctgt acctgaagggt 360
gaaggggaat gtgttcaaaa acaagcggat tctcatggaa cacatccaca agctgaaggc 420
agacaaggcc cgcaagaa 438

<210> 59
<211> 145
<212> PRT
<213> Tobacco mosaic virus

<400> 59
Lys Arg Leu Ala Ala Ser Val Met Lys Cys Gly Lys Gly Lys Val Trp
1 5 10 15
Leu Asp Pro Asn Glu Ser Ser Asp Ile Ser Met Ala Asn Ser Arg Gln
20 25 30
Asn Ile Arg Lys Leu Val Lys Asp Gly Phe Ile Ile Arg Lys Pro Thr
35 40 45
Lys Ile His Ser Arg Ser Arg Ala Arg Lys Met Lys Ile Ala Lys Met
50 55 60

Lys Gly Arg His Ser Gly Tyr Gly Lys Arg Lys Gly Thr Arg Glu Ala
 65 70 75 80
 Arg Leu Pro Thr Lys Val Leu Trp Met Arg Arg Met Arg Val Leu Arg
 85 90 95
 Arg Leu Leu Lys Lys Tyr Arg Glu Thr Lys Lys Ile Asp Lys His Met
 100 105 110
 Tyr His Asp Met Tyr Met Arg Val Lys Gly Asn Val Phe Lys Asn Lys
 115 120 125
 Arg Val Leu Met Glu Ser Ile His Lys Ser Lys Ala Lys Leu Gly Glu
 130 135 140
 Lys
 145

<210> 60
 <211> 147
 <212> PRT
 <213> Homo sapiens

<400> 60
 Lys Arg Leu Ala Ser Ser Val Leu Arg Cys Gly Lys Lys Lys Val Trp
 1 5 10 15
 Leu Asp Pro Asn Glu Thr Asn Glu Ile Ala Ala Asn Ala Asn Ser Arg
 20 25 30
 Gln Gln Ile Arg Lys Leu Ile Lys Asp Gly Leu Ile Ile Arg Lys Pro
 35 40 45
 Val Thr Val His Ser Arg Ala Arg Cys Arg Lys Asn Thr Leu Ala Arg
 50 55 60
 Arg Lys Gly Arg His Met Gly Ile Gly Lys Arg Lys Gly Thr Ala Asn
 65 70 75 80
 Ala Arg Met Pro Glu Lys Val Thr Trp Met Arg Arg Met Arg Ile Leu
 85 90 95
 Arg Arg Leu Leu Arg Arg Tyr Arg Glu Ser Lys Lys Ile Asp Arg His
 100 105 110
 Met Tyr His Ser Leu Tyr Leu Lys Val Lys Gly Asn Val Phe Lys Asn
 115 120 125
 Lys Arg Ile Leu Met Glu His Ile His Lys Leu Lys Ala Asp Lys Ala
 130 135 140
 Arg Lys Lys
 145

<210> 61
 <211> 11
 <212> DNA
 <213> Phage

<220>
 <221> promoter
 <222> (0)...(0)
 <223> T7 promoter

<400> 61
 tatagtattt t

11

<210> 62

<211> 11	
<212> DNA	
<213> Tobacco mosaic virus	
<400> 62	
tataggtatt t	11
<210> 63	
<211> 11	
<212> DNA	
<213> Tobacco mosaic virus	
<220>	
<221> promoter	
<222> (0)...(0)	
<223> N= Random base	
<400> 63	
tatagntatt t	11
<210> 64	
<211> 13	
<212> DNA	
<213> Tobacco mosaic virus	
<220>	
<221> promoter	
<222> (0)...(0)	
<223> N= A single random base	
<400> 64	
tatagtngta ttt	13
<210> 65	
<211> 12	
<212> DNA	
<213> Nicotiana benthamiana	
<400> 65	
tataggtatt tt	12
<210> 66	
<211> 14	
<212> DNA	
<213> Nicotiana benthamiana	
<400> 66	
tatagtcgta tttt	14
<210> 67	
<211> 15	
<212> DNA	
<213> Nicotiana benthamiana	

<220>
 <221> tRNA
 <222> (0)...(0)
 <223> N= A random base
 TMV vector RNA transcripts

<400> 67
 tatagtngtn gtatt 15

<210> 68
 <211> 21
 <212> DNA
 <213> Tobacco mosaic virus

<220>
 <223> N= A random base

<400> 68
 tatagtngtn gtngtngtat t 21

<210> 69
 <211> 16
 <212> DNA
 <213> Tobacco mosaic virus

<400> 69
 tatagtattt gtattt 16

<210> 70
 <211> 10
 <212> DNA
 <213> Viral

<220>
 <221> promoter
 <222> (0)...(0)
 <223> cDNA sequence

<400> 70
 tataggtatt 10

<210> 71
 <211> 12
 <212> DNA
 <213> Tobacco mosaic virus

<220>
 <221> promoter
 <222> (0)...(0)
 <223> GTC bases and virus cDNA

<400> 71
 tatagtcgta tt 12